Evaluation of Risk Factors for Nasopharyngeal Carcinoma in High-Risk Nasopharyngeal Carcinoma Families in Taiwan

Xiaohong (Rose) Yang,[1] Scott Diehl,[2] Ruth Pfeiffer,[1] Chien-Jen Chen,[3] Wan-Lun Hsu,[3] Mustafa Dosemeci,[1] Yu-Juen Cheng,[3] Brenda Sun,[4] Alisa M. Goldstein,[1] Allan Hildesheim,[1] and The Chinese and American Genetic Epidemiology of NPC Study Team

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS, Bethesda, Maryland; [2]Center for Pharmacogenomics and Complex Disease Research, New Jersey Dental School, University of Medicine and Dentistry of New Jersey, Newark, New Jersey; [3]Graduate Institute of Epidemiology, College of Public Health, National Taiwan University, Taipei, Taiwan; [4]Westat, Inc., Rockville, Maryland

Abstract

A study of nasopharyngeal carcinoma (NPC) families with two or more affected members was conducted in Taiwan (265 families with 2,444 individuals, 502 affected and 1,942 unaffected) to determine the association between NPC and potential etiologic factors in NPC high-risk families. Similar to results from a previous case-control study in Taiwan, Guangdong salted fish consumption during childhood, exposure to wood, and betel nut consumption were all associated with elevated NPC risk using conditional logistic regression, although these associations were not as strong as in the case-control study possibly due to shared environment among family members. Risk associated with cumulative wood exposure and salted fish consumption before age 10 was stronger in families with early NPC age-onset [odds ratio (OR_{wood}), 5.10; 95% confidence interval (95% CI), 1.50-17.34; OR_{fish}, 3.94; 95% CI, 1.47-10.55] or three or more affected members (OR_{wood}, 4.41; 95% CI, 1.58-12.30; OR_{fish}, 4.27; 95% CI, 1.10-16.47). In contrast, a tendency for elevated risk was noted for betel nut use in late age-onset families (OR, 2.44; 95% CI, 1.16-5.13) and the *CYP2E1 c2* allele in families with less than three affected members (OR, 2.06; 95% CI, 1.04-3.35). Risk estimates associated with these exposures were similar when the analyses were restricted to EBV-seropositive subjects. To better adjust for degree of relationship among family members and residual genetic correlations, we also calculated ORs using a variance components model. The results from the two methods were similar indicating that the risk estimates from conditional logistic regression were unbiased. (Cancer Epidemiol Biomarkers Prev 2005;14(4):900–5)

Introduction

Nasopharyngeal cancer (NPC) is a rare disease in most parts of the world where the incidence rate is generally <1 per 100,000 person-years. However, it is one of the most common tumors occurring in individuals in southern China and Southeast Asia, with the incidence of 20 to 40 per 100,000 person-years (1). Over the years, numerous studies have shown that the etiology of NPC is multifactorial, including genetic, environmental, and virological factors. Among these factors, EBV seems to be the most important. It has been well established that there is a strong association between infection of EBV and NPC (2). However, given the universality of EBV infection but the unique geographic distribution of NPC, factors other than EBV are also believed to be important determinants of the risk for NPC.

Familial clustering of NPC has been documented in Chinese populations (3, 4) and even in low-risk populations (5). Family history of NPC has been suggested to be associated with NPC risk (6). Several studies showed that NPC risk was significantly higher in first-degree relatives of cases than in the general population (6-8). These data suggested that genetic factors may also contribute to the development of NPC. Recent studies showed that genetic polymorphisms of some metabolic enzyme genes (*CYP2E1* and *GSTM1*; refs. 9-11) and some DNA repair genes (*XRCC1* and *hOGG1*; ref. 12) influenced susceptibility to NPC. Certain *HLA* groups were also reported to be associated with increased NPC risk (13). In addition, three genetic linkage studies carried out in families at high risk of NPC from southeast China, southern China, and Taiwan have mapped major susceptibility loci to chromosomes 4, 3, and 14, respectively (14, 15).[5]

Epidemiologic studies have suggested that a number of environmental factors may increase the risk of sporadic NPC; these factors include cigarette smoking, occupational exposures to wood and possibly, formaldehyde, and salted fish consumption (2, 16-21). In our previous evaluations from a case-control study conducted in Taiwan, we have shown that increased risks of NPC were associated with occupational exposure to wood dust for longer than 10 years (20); high intakes of nitrosamines and nitrite from meat, fish, and preserved vegetables during childhood and weaning (21); long-term (≥25 years) cigarette smoking (18); homozygosity for the *c2* allele of the *CYP2E1* gene (9); and several *HLA* alleles and haplotypes (13). The effects of such environmental and host factors in NPC high-risk families have not been characterized to date.

In this study, we examined environmental and some host exposures in multiplex families with more than two affected people collected in Taiwan and compared the NPC risk

Received 9/15/04; revised 11/16/04; accepted 12/14/04.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Xiaohong (Rose) Yang, Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Boulevard, Room 7014, Bethesda, MD 20852. Phone: 301-594-7804; Fax: 301-402-4489. E-mail: royang@mail.nih.gov

[5] S.R. Diehl et al. Genome-wide linkage and disequilibrium analysis of nasopharyngeal carcinoma demonstrate a complex etiology and identify multiple candidate genes. 2004, submitted for publication.

associated with these exposures to that of sporadic NPC. A better understanding of factors involved in the etiology of NPC is essential for counseling high-risk family members and for guiding future research of familial NPC kindreds.

Materials and Methods

Study Population. This study originated from a family study initiated in 1996 by NIH and National Taiwan University. The design of the family study was previously described.[5] Briefly, 475 potentially eligible families with two or more nonparent-offspring NPC subjects (in first-, second-, and third-degree relatives) were identified after an extensive review of records from the nationwide tumor registry in Taiwan, supplemented by information from listings obtained from 10 tertiary care hospitals throughout the Island and prospective identification at select outpatient clinics that treat NPC. In addition to recruiting all available affected family members, we also recruited up to five unaffected siblings and the parents of affected subjects. If NPC cases had adult children, we recruited their spouses and up to three unaffected offspring. When NPC cases were more distant relatives such as cousins, we recruited their shared second-degree relatives such as unaffected aunts and uncles to genetically connect the NPC cases. All subjects provided informed consent to participate, and the protocol was reviewed and approved by human subjects review committees in Taiwan and the United States.

Questionnaire. A risk factor questionnaire was given to all eligible individuals, as detailed previously (9). About half (n = 203, 42.5%) of the NPC cases were deceased at the time of interview and proxy interviews were conducted for these individuals with a close family member [spouses (38%), siblings (33%), and children (29%)]. Most unaffected family members (96.6%) were self-respondents. Risk factor interviews obtained information on demographic factors (age, sex, ethnicity, and education), cigarette smoking, betel nut and alcohol consumption, diet, and occupational exposures. A dietary questionnaire included questions about the consumption of Guangdong moldy salted fish, Guangdong firm salted fish, and other types of salted fish during childhood (before age 10) and as an adult (between age 10 and 30). Information on occupational history was obtained by reports of all jobs held for ≥1 year since the age of 16. The assessment of the occupational history data was detailed previously (20). Briefly, occupational history data were obtained from each participant and reviewed by an expert industrial hygienist. The intensity of exposure to wood, formaldehyde, and organic solvents was classified on a scale of 0 (not exposed) to 9 (strong) as intensity index for each exposure. For each subject, this information was combined with duration data to estimate the cumulative exposure (defined as duration of exposure × average intensity). In addition, duration of exposure was also calculated with the exclusion of exposures occurring in the 10 years preceding diagnosis (cases) or interview (controls).

Biological Specimen Collection and Testing. Each study participant was asked to consent to the collection of 30 mL blood for genotyping and EBV antibody testing. Serum was tested for antibody titers to several anti-EBV antibodies known to be associated with NPC, including viral capsid antigen IgA, EBV nuclear antigen 1 IgA, and anti-DNase antibodies, according to the methods described previously (22). Individuals positive for any of these EBV markers were considered seropositive and individuals negative to all were classified as seronegative. For *CYP2E1* genotyping, DNA extracted from peripheral blood mononuclear cells was tested for three polymorphisms of *CYP2E1* (*RSA1*, *PST1*, and *DRA1*) by PCR-based restriction fragment length polymorphism analysis described previously (9).

Data Analysis. Conditional logistic regression was used to obtain odds ratios (OR) and 95% confidence intervals (95% CI) between NPC risk and each exposure, matching on families. Several measures of substance (cigarette, alcohol, and betel nut) use and occupational exposures (wood, formaldehyde, and organic solvents) were analyzed separately [ever/never, total years of exposure, years of exposure before diagnosis (substance use), years of exposure excluding 10 years before the diagnosis/interview (occupational), age at first exposure, and cumulative exposure]. ORs obtained by using each individual variable resembled one another; thus, we selected years of consumption of substances before diagnosis and cumulative occupational exposures in final analyses. For diet analysis, adult consumption and childhood consumption of each food group were analyzed separately. We initially evaluated intake of both moldy and firm types of Guangdong salted fish separately. The numbers of subjects with childhood intake were small; thus, we combined both types of fish consumption. The three polymorphisms of *CYP2E1* (*RSA1*, *PST1*, and *DRA1*) were strongly correlated and thus only *RSA1* data was presented. The numbers of subjects with *c2-c2* genotype were small (3 affected and 31 unaffected); thus, we combined subjects with *c1-c2* and *c2-c2* genotype. Because age and sex were strong determinants of NPC risk, all analyses were adjusted for age (<40, 40-55, and >55 years) and sex. Variables in the final multivariate model also included cigarette smoking, betel nut use, Guangdong salted fish consumption during childhood, and cumulative wood exposure. Additional adjustment for education, ethnicity, other salted fish consumption during childhood and cumulative formaldehyde exposure did not affect risk estimates and thus were removed from the final multivariate model. *RSA* genotypic data was available for only a subset (27%) of all subjects recruited earlier in the study, and including *RSA* in the multivariate model did not change the risk estimates for other variables. Therefore, *RSA* variable was also removed from the final model and presented separately.

Because the family study is still ongoing, occupational data were not available for the most recently recruited subjects (40%). We believe these variables were missing completely at random because the availability of data was determined by the time of recruitment and should not introduce bias in exposure levels. In addition, 10% of the recruited subjects responded ''don't know'' to the question about the consumption of salted fish. Missing data for these variables were accounted for in the multivariable model as dummy categories. To avoid the potential bias caused by this approach, we also analyzed the multivariable model for subjects with complete data only. Risk estimates for all variables did not change measurably and thus we included the missing categories in the final model.

Because EBV is a necessary risk factor for NPC and NPC patients consistently show higher antibody titers to various EBV-associated antigens before treatment compared with controls (1), we considered all NPC cases as EBV positives. The effect of EBV seropositivity on risk of NPC associated with other exposures was assessed by analyses restricted to all cases (n = 502) and EBV seropositive controls (n = 501).

To reduce possible reporting bias by surrogate respondents, we also repeated all analyses restricted to self-respondents. Because the results were essentially unchanged, we combined self-respondents and surrogate respondents in the subsequent analyses. We also evaluated risk according to degree of relationship of controls to NPC cases. There were only 29 (1.5%) unaffected individuals who did not have a first-degree affected relative and 172 (8.9%) who were spouses of a NPC case. ORs between NPC risk and all variables calculated among first-degree relatives were quite similar to those obtained from all individuals. Therefore, we included all

subjects in the final model. All conditional logistic regression analyses were done using the PROC PHREG procedure included in the SAS (version 8.0, SAS Institute, Inc., Cary, NC) software.

Risk estimates associated with cigarette smoking, betel nut use, occupational exposures to wood and formaldehyde, Guangdong fish consumption during childhood, and *RSA* at risk allele were also calculated by comparing the familial cases to community controls collected from our previously conducted case-control study (9), using unconditional logistic regression (PROC LOGISTIC in SAS).

Previous work indicated that for rare diseases like NPC with a multifactorial/polygenic etiology, use of conditional logistic regression models that condition on family will yield largely unbiased estimates of environmental risk factors if residual genetic correlations among family members are modest (23). To assess if the presented estimates are unbiased, the final choice of variables was also analyzed in a variance components model proposed by Pfeiffer et al. (24), which accounts for different residual correlations among family members. Risk estimates from the variance components model for each variable were compared with those obtained from conditional logistic regression. All statistical tests were two sided.

Results

Risk factor questionnaire data was available from 265 families consisting of a total of 2,444 (502 affected and 1,942 unaffected) individuals. The distribution of the number of affected and unaffected individuals within these families is shown in Table 1. Demographic (age, sex, ethnicity, and education) and substance (cigarette smoking, betel use, and alcohol drinking) consumption data were available for most (>99%) respondents. Occupational and dietary data were available for 59% (65% cases and 57% controls) and 89% (82% cases and 91% controls) subjects, respectively. EBV testing results were available for 1,771 unaffected family members (91.1%). All NPC cases were considered EBV positive. Genotyping results for *RSA1* of *CYP2E1* were available on 103 affected (21%) and 553 (28%) unaffected individuals.

NPC cases were older (mean age = 51.6 years) than controls (mean age = 46.5 years) at the time of the interview and more often male (71.7%) among NPC cases compared with controls (47.2%). In addition, cases were less educated than controls; 41.9% of cases finished senior high school or higher education versus 52.7% of controls. Cases and controls were similar with regard to ethnicity, diagnosis of other types of cancer, consumption of alcohol and salted fish (moldy and firm) after age 10, and occupational exposure to formaldehyde and solvents (data not shown). In age- and sex-adjusted analysis, the measured variables were not associated with statistically significant elevations in NPC risk, although OR for heavy wood exposure (≥25) was suggestive (OR, 1.90; 95% CI, 0.88-4.13). Results of the multivariable regression analysis of multiple risk factors of NPC are shown in Table 2. The adjusted OR for cumulative wood exposure (≥25) was significant; the ORs for longer betel use (≥20 years) and frequent intake of Guangdong salted fish (≥1/wk) during childhood were suggestive but not significant (Table 2). Interestingly, cigarette smoking, especially with longer duration (≥20 years), seemed inversely related to NPC risk. The patterns observed were similar in analyses restricted to all cases and 501 controls seropositive for antibodies against EBV, except for the attenuated effect of *RSA* allele (Table 2).

Table 1. Distribution of affected and unaffected individuals in NPC high-risk families (n = 265)

	Total no. families
No. NPC subjects in family*	
1	69
2	155
3-5	39
>5	2
No. unaffected in family*	
≤5	107
6-10	93
>10	74

*These values reflect number of recruited cases within a family; All families have two or more affected individuals; however, most families had two or less cases recruited in this study.

To assess if the presented estimates are unbiased, the final multivariate model was also analyzed in a variance components model. Results from the variance components model were very similar to the OR estimates obtained from conditional models, with the changes of ORs all under 10% (range, ±0.05% to 7.6%). We therefore used conditional logistic regression in all subsequent stratified analyses.

To further determine whether the risk associated with exposure variables was modified by genetic background, we evaluated the subset of families with a higher potential for genetic risk of NPC. We did stratified analyses based on the numbers of affected NPC cases and age-onset of NPC within a family. Families with ≥3 affected cases or families with at least one affected individual diagnosed before age 40 were defined as families with high genetic risk. The risk of NPC associated with heavy wood exposure (≥25) and frequent Guangdong salted fish intake (≥1/wk) during childhood was elevated in families with high genetic risk (Table 2). The test for interaction between genetic risk (≥3 affected cases) and heavy wood exposure was statistically significant (P = 0.018). An interaction between genetic risk (≥3 affected cases) and frequent salted fish intake was not statistically significant (P = 0.23), possibly due to the small number in the exposed group. Betel nut use and *RSA* allele, on the other hand, seemed to have stronger effects in families with low genetic risk (Table 2). Long duration of cigarette smoking (≥20 years) showed a consistently inversely related effect in all stratified analyses (Table 2).

To examine the possible effect of shared environment among affected and unaffected family members on the risk estimates, we compared NPC risk associated with these exposures between familial cases (n = 502) and community controls (n = 327) ascertained in the previously conducted case-control study in Taiwan (9). Table 3 shows the risk estimates obtained from the comparisons of sporadic cases against community controls, familial cases against family-based controls, and familial cases against community controls, respectively. Compared with the data from the community-based case-control study, the risk of NPC associated with each of these exposures was attenuated in the family-based study. However, when comparing familial cases with the community controls, age- and sex-adjusted ORs associated with betel nut use, formaldehyde exposure, and Guangdong salted fish consumption during childhood showed stronger effects, with ORs even higher than those obtained from the community-based case-control comparisons. Risks associated with smoking, *RSA c2* allele, and wood exposure did not change significantly regardless of which control group was used. To match the geographic location of the community controls selected in the Taipei area (9), we restricted the familial cases to Taipei residents and the results were not measurably changed (data not shown).

Discussion

In this study, we evaluated some potential NPC risk factors in NPC multiplex families with two or more affected individuals collected in Taiwan and compared the risk associated with

Table 2. Adjusted ORs in stratified analyses

	All subjects (n = 2,442, affected 502, unaffected 1,942)			EBV-positive (n = 1003, affected 502, unaffected 501)	Early age-onset (<40 y)*, (n = 1,236, affected 257, unaffected 979)	Late age-onset (≥40 y), (n = 1,206, affected 227, unaffected 963)	Families with higher no. cases (>3)†, (n = 841, affected 181, unaffected 660)	Families with lower no. cases (≤2), (n = 1,603, affected 321, unaffected 1,282)
Variables	Case (%)	Control (%)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Duration of smoking (y)‡								
Never	269 (58.3)	1,328 (68.3)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)
<20	100 (21.7)	385 (19.8)	0.80 (0.57-1.13)	0.59 (0.35-0.99)	0.57 (0.36-0.91)	1.41 (0.77-2.56)	0.60 (0.34-1.05)	0.90 (0.59-1.39)
≥20	92 (20.0)	232 (11.9)	0.66 (0.45-0.96)	0.58 (0.34-1.00)	0.57 (0.32-1.00)	0.78 (0.45-1.35)	0.56 (0.29-1.05)	0.67 (0.42-1.06)
Duration of betel nut (y)‡								
Never	429 (87.0)	1,759 (90.0)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)
<20	27 (5.5)	137 (7.0)	0.77 (0.46-1.28)	0.62 (0.30-1.25)	1.08 (0.56-2.06)	0.62 (0.25-1.52)	0.77 (0.35-1.70)	0.70 (0.36-1.35)
≥20	37 (7.5)	59 (3.0)	1.69 (0.97-2.92)	1.36 (0.65-2.89)	0.71 (0.27-1.84)	2.44 (1.16-5.13)	1.38 (0.50-3.83)	1.67 (0.89-3.13)
Cumulative wood§								
None	303 (93.2)	1,062 (94.9)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)
<25	7 (2.2)	38 (3.4)	0.54 (0.22-1.36)	0.74 (0.21-2.62)	0.29 (0.16-0.54)	0.70 (0.20-2.49)	0.33 (0.06-1.65)	0.70 (0.23-2.17)
≥25	13 (4.6)	19 (1.7)	2.29 (1.04-5.07)	2.98 (0.93-9.51)	5.10 (1.50-17.34)	1.55 (0.47-5.08)	4.41‖ (1.58-12.30)	0.93 (0.23-3.75)
Guangdong salted fish (age <10)§								
Never	342 (93.4)	1,528 (93.4)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)
<1/wk	13 (3.6)	72 (4.4)	0.65 (0.31-1.35)	0.64 (0.23-1.74)	0.63 (0.26-1.57)	0.32 (0.07-1.51)	0.49 (0.13-1.81)	0.73 (0.29-1.80)
≥1/wk	11 (3.0)	36 (2.2)	1.78 (0.82-3.89)	1.74 (0.58-5.20)	3.94 (1.47-10.55)	0.67 (0.14-3.31)	4.27‖ (1.10-16.47)	1.28 (0.48-3.39)
RSA¶								
c1-c1	57 (55.3)	331 (59.9)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)
c1/2-c2	46 (44.7)	222 (40.1)	1.45 (0.79-2.65)	0.81 (0.35-1.85)	0.81 (0.43-1.54)	1.62 (0.47-5.62)	0.42 (0.15-1.21)	2.06 (1.04-3.35)

*At least one affected individual in that family had early age-onset.

† No. cases in a family was based on the total number of affected subjects, regardless of how many were recruited in that family.

‡ Total years of exposure before diagnosis.

§Occupational and dietary data were available for 59% (65% cases and 57% controls) and 89% (82% cases and 91% controls) subjects, respectively.

‖The interaction was statistically significant between genetic risk (defined as having three or more cases within a family) and heavy wood exposure (P = 0.018), but was not significant with frequent fish consumption (P = 0.23). Interactions were tested by including an interaction term in the model.

¶Combined *c1-c2* and *c2-c2* genotype. Genotyping results were available on 103 affected (21%) and 553 (28%) unaffected individuals. Adjusted for age, sex, cigarette smoking, betel nut consumption, wood and formaldehyde exposure, and Guangdong and other salted fish consumption during childhood.

these exposures to that of sporadic NPC estimated from a previously reported case-control study also conducted in Taiwan. Overall, a similar pattern of associations was observed in the two studies, with the exception of cigarette smoking. Similar to the results obtained from the case-control study, Guangdong salted fish consumption during childhood (≥1/wk), exposure to wood (≥25), betel nut consumption (≥20 years), and possibly, having a *CYP2E1 RSA c2* allele, were all associated with elevated NPC risk, although these associations were not as strong as in the case-control study. However, unlike the results from the case-control study, smoking seemed inversely related to NPC risk in these multiplex families, especially among smokers who had smoked for >20 years before their NPC diagnosis.

To our knowledge, our study is the first to examine NPC risk factors among members of NPC high-risk families. Family history of NPC has been suggested as a risk factor for NPC (3, 4), but the mechanism of familial clustering is not understood. It may reflect genetic factors, shared environmental factors, or both. The results seen in our high-risk families suggest that members from families with putatively higher genetic risk (early age-onset or ≥3 affected people within a family) may be at greatly elevated risk of NPC if they are heavily exposed to wood or frequently eat Guangdong salted fish during childhood. The point estimates of both exposures in families with high genetic risk were much higher compared with those obtained from families with less affected people and late age-onset; they were even higher than those obtained from the case-control study. These results may reflect interactions between these exposures and genetic predisposition. In contrast, betel nut use with long duration and *RSA* at risk allele had stronger effects in families with lower genetic risk. Because the risk associated with *RSA* polymorphism in families with lower genetic risk is similar to the risk in the general population, our result is consistent with the previous report that the *RSA* at risk allele is a risk factor for NPC. The lower risk in families with high genetic risk may reflect the higher frequency of this polymorphism in these families, even among unaffected people. These findings suggest that different environmental exposures may affect, to varying degrees and with varying directions, the risk of NPC in familial NPC kindreds. However, these results have to be interpreted with caution because of the small number in heavy exposure groups. Further research is needed to examine interactions between these exposures and genetic predisposition.

The etiology of NPC has been indicated by numerous studies to be multifactorial, involving both genetic and environmental factors. More recently, Jia et al. did a segregation analysis of a total of 1,903 Cantonese pedigrees ascertained in Guangzhou, China. The results of that study supported a multifactorial mode of inheritance, with no evidence for a substantial role of major genes in the inheritance of NPC (25). These data indicate that shared environmental exposures among family members may be a significant contributor to the familial clustering of NPC. Most of the risk factors measured in our study seemed to have a weaker effect in familial NPC compared with sporadic NPC. To further examine whether this was due to the shared environment among family members, we compared the risk estimates obtained from the family-based case-control comparisons with those observed from familial case-community control comparisons that were measured from the previously conducted case-control study (9, 13, 18, 20, 21). Compared with family-based case-control comparisons and community-based case-control comparisons, familial case against population control comparisons showed stronger effects associated with

the long duration of betel nut use, frequent intake of Guangdong salted fish during childhood, and heavy formaldehyde exposures (Table 3), indicating that the attenuation of risk estimates with familial case to family control comparisons was at least partly due to the shared environment among affected and unaffected family members.

Results from the epidemiologic studies examining the association between cigarette smoking and the risk of NPC have been inconsistent. Whereas some studies suggested cigarette smoking to be a risk factor for NPC (17, 18, 26, 27), others failed to observe this association (28-31). In our previously reported case-control study conducted in Taiwan, long duration of smoking before diagnosis was significantly associated with NPC (18). In contrast, in our family study, smoking before diagnosis seemed inversely related to NPC risk, especially for long duration (≥25 years). Similar results were obtained from analyses using ever/never smoking status, smoking intensity, and pack-years of smoking (data not shown). In addition, in contrast to an earlier report that age at start of smoking was inversely associated with the risk of NPC (32), our data showed no such association ($OR_{age\ 18\text{-}21}$, 0.92; 95% CI, 0.66-1.28; $OR_{age\ >22}$, 1.03; 95% CI, 0.75-1.42). The inversely related effect of cigarette smoking was even stronger in familial case to population control comparisons, suggesting that familial cases smoked less compared with either control groups. Because data from about half of the familial cases were collected from proxy interviews, we restricted our analyses to self-respondents to avoid the potential bias from surrogate report. Results from self-respondent data were generally consistent with those obtained from all respondents, suggesting that this is unlikely a strong source of bias in our study. However, because some of these familial cases were not incident cases and there was a delay between time of diagnosis and data collection, cases may have been more aware of the risk factors for NPC, such as smoking, than controls. This potential bias may have, at least partly, explained why familial cases tended to report smoking less. Because this is the first effort to examine the effect of nongenetic factors on familial NPC risk, this is an interesting finding that needs to be further evaluated.

We used conditional logistic regression to account for some dependency within families. However, this approach can sometimes lead to underestimates of exposure effects if genetic correlations are ignored (24). Pfeiffer et al. proposed a variance components or random effects model to further account for common familial effects and for different genetic correlations among family members (24). Subsequently, they investigated the robustness of this model with respect to various misspecifications of genetic random effects in simulations. Their results showed that if the disease of interest is rare and includes an unmeasured genetic component that is modest, then conditional logistic regression yields nearly unbiased estimates (23). In our study, we also analyzed the NPC family data using this variance components model and showed that the risk estimates were similar to those obtained from conditional logistic regression. Therefore, conditional logistic regression provides a reasonable analytic strategy for our study.

It should be noted that our study has several limitations. As we mentioned above, information on about half of the cases was obtained from proxy respondents, because many cases were deceased at the time of interview. To evaluate whether proxy

Table 3. Risk estimates from case-control and family studies

Variables	Case-control study (n_{case}=325, $n_{control}$=327), OR* (95% CI)	Family study	
		Familial cases to family controls (n_{case}=502, $n_{control}$=1,944)	Familial cases to population controls (n_{case}=502, $n_{control}$=327)
		OR† (95% CI)	OR‡ (95% CI)
Duration of smoking (y)			
Never	1 (reference)	1 (reference)	1 (reference)
<25	1.20 (0.80-2.00)	0.88 (0.63-1.22)	0.74 (0.49-1.11)
≥25	1.70 (1.09-2.74)	0.75 (0.53-1.07)	0.60 (0.39-0.93)
Duration of betel nut use (y)			
Never	1 (reference)	1 (reference)	1 (reference)
<20	1.08 (0.57-2.07)	0.68 (0.42-1.10)	1.22 (0.65-2.30)
≥20	1.37 (0.61-3.10)	1.36 (0.81-2.27)	1.83 (1.08-4.67)
Cumulative wood exposure			
None	1 (reference)	1 (reference)	1 (reference)
<25	1.20 (0.58-2.50)	0.49 (0.20-1.19)	0.32 (0.13-0.77)
≥25	2.40 (1.20-5.10)	1.90 (0.88-4.13)	1.74 (0.69-4.39)
Cumulative formaldehyde exposure			
None	1 (reference)	1 (reference)	1 (reference)
<25	1.30 (0.70-2.40)	1.03 (0.60-1.76)	1.30 (0.70-2.39)
≥25	1.50 (0.88-2.70)	1.31 (0.87-1.97)	4.29 (2.45-7.51)
Guangdong salted fish consumption (age <10)			
Never	1 (reference)	1 (reference)	1 (reference)
Ever	1.40 (0.47-4.50)	0.92 (0.54-1.57)	4.73 (1.75-12.78)
RSA§			
c1-c1	1 (reference)	1 (reference)	1 (reference)
c1-c2	0.83 (0.60-1.20)	1.44 (0.82-2.54)	1.34 (0.84-2.14)
c2-c2	2.60 (1.20-5.74)	0.68 (0.16-2.86)	1.05 (0.27-4.12)

*ORs obtained from the case-control study for duration of smoking before diagnosis, cumulative wood and formaldehyde exposure, Guangdong salted fish consumption before age 10, and *RSA* allele were previously published (9, 18, 20, 21). OR for the duration of betel nut use before diagnosis was calculated using data collected in the same case-control study using unconditional logistic regression. ORs for smoking, betel nut use, Guangdong salted fish consumption, and *RSA* allele carrier were adjusted for age and sex. ORs for wood and formaldehyde exposure were adjusted for age, sex, education, and ethnicity.

† All ORs were calculated using conditional logistic regression and were adjusted for age and sex; Guangdong salted fish consumption was presented as ever/never to compare with that of the case-control study.

‡ All ORs were calculated using unconditional logistic regression and were adjusted for age and sex.

§To compare with the published result from the case-control study, *RSA c1-c2* and *c2-c2* genotype were analyzed separately; however, because only three cases in the family study had *c2-c2*, the risk estimate in this category from the conditional logistic regression might not be accurate.

data biased our findings, we did separate analyses among self-respondents. The findings of our study were unchanged when the analyses were restricted to the self-respondents indicating that inclusion of proxy data is unlikely to have introduced a strong source of bias. The distributions of cases in each exposure category obtained from self-respondents and all respondents were quite similar, with a maximal difference of 1.7%. Differences in ORs obtained from the two analyses were also small (0.56% for frequent salted fish consumption during childhood, 17% for heavy wood exposure, 9% for long duration of smoking before diagnosis, and 3.6% for long duration of betel nut use). Of note, when compared with the ORs obtained from self-respondents, the direction of association for the risk factors was invariably towards the null. Another limitation of our study is that familial cases, and/or surrogates for cases, might have been more aware of some risk factors for NPC than controls, which might have caused a reporting bias. Finally, because we were dealing with family data, we had to face the inevitable challenge of missing data. We included the missing data in the multivariate model as dummy categories to avoid unnecessarily discard of data. To avoid the potential bias caused by this approach, we also analyzed the multivariable model for subjects with complete data only. Except for the fish variable, for which the change of OR was nearly 30%, the risk estimates for all other variables did not change measurably (range, 0-0.8%). The more noticeable change for the fish variable was possibly due to the extremely small number in the heavy exposure group when using the complete data.

In conclusion, the results from this study suggested that NPC risk factors, such as salted fish consumption, wood exposure, betel nut use, and possibly, *CYP2E1* polymorphism, might be associated with an increased risk of familial NPC. People at putatively higher genetic risk might be at further increased risk if they have some of these exposures. In contrast to what has been suggested in some population-based studies, smoking was not associated with increased NPC risk in our family study. Overall, our results confirmed the risk of these exposures and emphasized the need for additional research focusing on the interplay of genetic and environmental factors in the etiology of NPC.

Appendix 1

The Chinese and American Genetic Epidemiology of NPC Study Team includes Czau-Siung Yang and Jen-Yang Chen (Department of Microbiology, National Taiwan University), Mei-Ying Liu (National Taipei College of Nursing), Mow-Ming Hsu and Tzung-Shian Sheen (Department of Otolaryngology, Chung-Shan Medical and Dental College Hospital), Ming-Tsai (Department of Otolaryngology, China Medical College Hospital), Sen-Tien Tsai and Wen-Yuan Chao (Department of Otolaryngology, National Cheng-Kung University Hospital), Peir-Rong Chen (Department of Otolaryngology, Buddhist Tzu-Chi General Hospital), Ih-Jen Su and Hey-Chi Hsu (Department of Pathology, National Taiwan University Hospital), Chi-Long Chen (Department of Pathology, China Medical College Hospital), Ching-Kuang Tai (Gee-Tien Ear Nose and Throat Hospital), Luo-Ping Ger (Department of Medical Education and Research, Kaohsiung Veteran General Hospital), and I-How Chen and Ji-Hong Hong (Taipei Chang-Gung Memorial Hospital).

References

1. Yu MC, Henderson BE. Nasopharyngeal cancer. In: Schottenfeld D, Fraumeni JF Jr, editors. Cancer epidemiology and prevention. New York: Oxford University Press; 1996. p. 603–18.
2. Hildesheim A, Levine PH. Etiology of nasopharyngeal carcinoma: a review. epidemiologic reviews 1993;15:466–85.
3. Ko JY, Sheen TS, Hsu MM, Lui LT. Familial clustering of nasopharyngeal carcinoma. Otolaryngol Head Neck Surg 1998;118:736–7.
4. Zhang F, Zhang J. Clinical hereditary characteristics in nasopharyngeal carcinoma through Ye-Liang's family cluster. Chin Med J(Engl) 1999;112: 185–7.
5. Coffin CM, Rich SS, Dehner LP. Familial aggregation of nasopharyngeal carcinoma and other malignancies. A clinicopathologic description. Cancer 1991;68:1323–8.
6. Chen CJ, Liang KY, Chang YS, et al. Multiple risk factors of nasopharyngeal carcinoma: Epstein-Barr virus, malarial infection, cigarette smoking and familial tendency. Anticancer Res 1990;10:547–53.
7. Ung A, Chen CJ, Levine PH, et al. Familial and sporadic cases of nasopharyngeal carcinoma in Taiwan. Anticancer Res 1999;19:661–5.
8. Zheng X, Yan L, Nilsson B, Eklund G, Drettner B. Epstein-Barr virus infection, salted fish and nasopharyngeal carcinoma. A case-control study in southern China. Acta Oncol 1994;33:867–72.
9. Hildesheim A, Anderson LM, Chen CJ, et al. CYP2E1 genetic polymorphisms and risk of nasopharyngeal carcinoma in Taiwan. J Natl Cancer Inst 1997;89:1207–12.
10. Kongruttanachok N, Sukdikul S, Setavarin S, et al. Cytochrome *P*450 2E1 polymorphism and nasopharyngeal carcinoma development in Thailand: a correlative study. BMC Cancer 2001;1:4.
11. Nazar-Stewart V, Vaughan TL, Burt RD, Chen C, Berwick M, Swanson GM. Glutathione *S*-transferase M1 and susceptibility to nasopharyngeal carcinoma. Cancer Epidemiol Biomarkers Prev 1999;8:547–51.
12. Cho EY, Hildesheim A, Chen CJ, et al. Nasopharyngeal carcinoma and genetic polymorphisms of DNA repair enzymes XRCC1 and hOGG1. Cancer Epidemiol Biomarkers Prev 2003;12:1100–4.
13. Hildesheim A, Apple RJ, Chen CJ, et al. Association of HLA class I and II alleles and extended haplotypes with nasopharyngeal carcinoma in Taiwan. J Natl Cancer Inst 2002;94:1780–9.
14. Feng BJ, Huang W, Shugart YY, et al. Genome-wide scan for familial nasopharyngeal carcinoma reveals evidence of linkage to chromosome 4. Nat Genet 2002;31:395–9.
15. Xiong W, Zeng ZY, Xia JH, et al. A susceptibility locus at chromosome 3p21 linked to familial nasopharyngeal carcinoma. Cancer Res 2004;64:1972–4.
16. Yu MC. Nasopharyngeal carcinoma: epidemiology and dietary factors. IARC Sci Publ 1991;105:39–47.
17. Nam JM, McLaughlin JK, Blot WJ. Cigarette smoking, alcohol, and nasopharyngeal carcinoma: a case-control study among U.S. Whites. J Natl Cancer Inst 1992;84:619–22.
18. Cheng YJ, Hildesheim A, Hsu MM, et al. Cigarette smoking, alcohol consumption and risk of nasopharyngeal carcinoma in Taiwan. Cancer Causes Control 1999;10:201–7.
19. Vaughan TL, Stewart PA, Teschke K, et al. Occupational exposure to formaldehyde and wood dust and nasopharyngeal carcinoma. Occup Environ Med 2000;57:376–84.
20. Hildesheim A, Dosemeci M, Chan CC, et al. Occupational exposure to wood, formaldehyde, and solvents and risk of nasopharyngeal carcinoma. Cancer Epidemiol Biomarkers Prev 2001;10:1145–53.
21. Ward MH, Pan WH, Cheng YJ, et al. Dietary exposure to nitrite and nitrosamines and risk of nasopharyngeal carcinoma in Taiwan. Int J Cancer 2000;86:603–9.
22. Pickard A, Chen CJ, Diehl SR, et al. Epstein-Barr virus seroreactivity among unaffected individuals within high-risk nasopharyngeal carcinoma families in Taiwan. Int J Cancer 2004;111:117–23.
23. Pfeiffer RM, Hildesheim A, Gail MH, et al. Robustness of inference on measured covariates to misspecification of genetic random effects in family studies. Genet Epidemiol 2003;24:14–23.
24. Pfeiffer RM, Gail MH, Pee D. Inference for covariates that accounts for ascertainment and random genetic effects in family studies. Biometrika 2001; 88:933–48.
25. Jia WH, Collins A, Zeng YX, et al. Complex segregation analysis of nasopharyngeal carcinoma in Guangdong, China: evidence for a multifactorial mode of inheritance (complex segregation analysis of NPC in China). Eur J Hum Genet 2004;13 Oct [Epub ahead of print].
26. Chow WH, McLaughlin JK, Hrubec Z, Nam JM, Blot WJ. Tobacco use and nasopharyngeal carcinoma in a cohort of US veterans. Int J Cancer 1993; 55:538–40.
27. Yuan JM, Wang XL, Xiang YB, Gao YT, Ross RK, Yu MC. Non-dietary risk factors for nasopharyngeal carcinoma in Shanghai, China. Int J Cancer 2000;85:364–9.
28. Henderson BE, Louie E, SooHoo JJ, Buell P, Gardner MB. Risk factors associated with nasopharyngeal carcinoma. N Engl J Med 1976;295:1101–6.
29. Sriamporn S, Vatanasapt V, Pisani P, Yongchaiyudha S, Rungpitarangsri V. Environmental risk factors for nasopharyngeal carcinoma: a case-control study in northeastern Thailand. Cancer Epidemiol Biomarkers Prev 1992; 1:345–8.
30. Lee HP, Gourley L, Duffy SW, Esteve J, Lee J, Day NE. Preserved foods and nasopharyngeal carcinoma: a case-control study among Singapore Chinese. Int J Cancer 1994;59:585–90.
31. Zou J, Sun Q, Akiba S, et al. A case-control study of nasopharyngeal carcinoma in the high background radiation areas of Yangjiang, China. J Radiat Res (Tokyo) 2000;41 Suppl:53–62.
32. Zhu K, Levine RS, Brann EA, Gnepp DR, Baum MK. Cigarette smoking and nasopharyngeal cancer: an analysis of the relationship according to age at starting smoking and age at diagnosis. J Epidemiol 1997;7:107–11.